

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-085183

(43)Date of publication of application : 30.03.1999

(51)Int.Cl. G10L 3/00
G10L 3/00

(21)Application number : 09-238565

(71)Applicant : ATR ONSEI HONYAKU TSUSHIN
KENKYUSHO:KK

(22)Date of filing : 03.09.1997

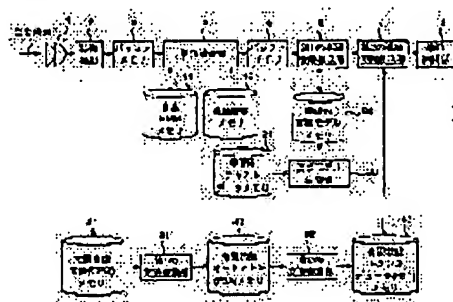
(72)Inventor : TSUKADA HAJIME
YAMAMOTO HIROSHI
KOSAKA YOSHINORI

(54) SPEECH RECOGNITION SYSTEM AND ITS APPARATUS AS WELL AS STORAGE MEDIUM RECORDING SPEECH RECOGNITION PROCESSING PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To recognize grammatical natural utterance with high accuracy and to discover a highly reliable section by combining statistical language restrictions and grammatical restrictions.

SOLUTION: The context free grammar(CFG) stored in a context free grammar memory 41 is previously subjected to approximate conversion to a finite state automata(FSA) by a first grammar conversion section 31 and is stored in a finite state automata memory 42. The stored finite state automata is stored into a finite state transducer(FST) memory 43 including the state transition of the addition, deletion and permutation of words by a second grammar conversion section 32. A second word temporary establishment pinpointing section 7 pinpoints the temporary establishment of the word string outputted from a first word temporary establishment pinpointing section 6 by using the finite state transducer(FST) and puts marks to the words deviating from the grammar. Finally a word extraction section 8 removes the marked words, extracts the utterance fragments consisting of the highly reliable word strings and outputs the same as speech recognition results.



LEGAL STATUS

[Date of request for examination] 03.09.1997

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3027557

[Date of registration] 28.01.2000

[Number of appeal against examiner's decision of

rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第3027557号
(P3027557)

(45) 発行日 平成12年4月4日(2000.4.4)

(24) 登録日 平成12年1月28日(2000.1.28)

(51) Int.Cl. ⁷	識別記号	FI	
G10L 15/18		G10L 3/00	537E
15/14			535Z

請求項の数 3 (全 15 頁)

(21) 出願番号	特願平9-238565	(73) 特許権者	593118597 株式会社エイ・ティ・アール音声翻訳通 信研究所 京都府相楽郡精華町大字乾谷小字三平谷 5番地
(22) 出願日	平成9年9月3日(1997.9.3)	(72) 発明者	塚田 元 京都府相楽郡精華町大字乾谷小字三平谷 5番地 株式会社エイ・ティ・アール音 声翻訳通信研究所内
(65) 公開番号	特開平11-85183	(72) 発明者	山本 博史 京都府相楽郡精華町大字乾谷小字三平谷 5番地 株式会社エイ・ティ・アール音 声翻訳通信研究所内
(43) 公開日	平成11年3月30日(1999.3.30)	(74) 代理人	100062144 弁理士 青山 稔 (外2名)
審査請求日	平成9年9月3日(1997.9.3)	審査官	涌井 智則

最終頁に続く

(54) 【発明の名称】 音声認識方法及び装置、並びに音声認識処理プログラムを記録した記録媒体

1

(57) 【特許請求の範囲】

【請求項1】 入力される発声音声文の音声信号に基づいて上記発声音声文を音声認識する音声認識方法であつて、

所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換するステップと、

上記変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換するステップと、

入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声文を音声認識するステップと、

2

上記音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力するステップとを含むことを特徴とする音声認識方法。

【請求項2】 入力される発声音声文の音声信号に基づいて上記発声音声文を音声認識する音声認識装置であつて、

所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換する第1の変換手段と、

上記第1の変換手段によって変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換する第2の変換手段と、

入力される発声音声文の音声信号に基づいて、所定の統

10

(2)

特許第3027557号

3

4

計的言語モデルと、上記第2の変換手段によって変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声を音声認識する音声認識手段と、

上記音声認識手段によって音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力する単語抽出手段とを備えたことを特徴とする音声認識装置、

【請求項3】 入力される発声音声文の音声信号に基づいて上記発声音声を音声認識する音声認識処理プログラムを記録した記録媒体であって、所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換するステップと、

上記変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換するステップと、

入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声を音声認識するステップと、

上記音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力するステップとを含む音声認識処理プログラムを記録したことを特徴とする音声認識処理プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声認識方法及び装置、並びに音声認識処理プログラムを記録した記録媒体に関する。

【0002】

【従来の技術】自然発話音声の認識においては、N-gramを基礎とする統計的言語モデルが広く使用されている。これは、こうしたモデルが、逸脱した発話を受容すると同時に探索の際の認識候補数を大幅に低減できるためである。一方で、音声翻訳装置を含む多くの音声対話システムでは、音声認識装置の後端部が文法を使用して構文構造を解析している。通常これは、音声認識に使用されるN-gramを基礎とする統計的言語モデルとは無関係に開発されている。N-gramを基礎とする統計的言語モデルと文法とは異なるタイプの言語制約として機能するため、音声対話システムの全体的性能を高めるためには、統計的言語モデルだけでなく後端部の文法をも音声認識上の制約として協働的に使用する必要が

ある。

【0003】統計的言語モデルと厳格な文法的制約の両方を統合する方法としては、文法を逸脱していない発話のみを受容するものが多く提案されている（例えば、特願平8-330679号など参照）。さらに、後置の文法によって近似された厳格な文法的制約を使用する認識方法が提案されている（例えば、従来技術文献「F.C. N. Pereira, et al., "Finite-state approximation of phrase-structure grammars", In 29th Annual Meeting of the Association for Computational Linguistics, pp.246-255, 1991年」参照。）、

【0004】

【発明が解決しようとする課題】しかしながら、こうした研究のように文法を自然発話音声認識に於ける厳格な制約として使用することには幾つかの欠点がある。第一に、自然発話音声は、言いどみ、言い直しといった朗読音声の場合には発生しない自然発話音声固有の言語的現象によって文法を外れることが多い。第二に多くの場合、厳格に文法的制約を適用することは頑強ではない。文法によって長期依存性を表現することは可能であるが、局部的エラーによって全体に悪影響が及ぶ場合が多い。

【0005】本発明の目的は以上の問題点を解決し、自然発話音声とその固有の言語的現象によって文法を外れた場合であっても、従来技術に比較して高い認識率で音声認識することができる音声認識方法及び装置、並びに、音声認識処理プログラムを記録した記録媒体を提供することにある。

【0006】

【課題を解決するための手段】本発明に係る音声認識方法は、入力される発声音声文の音声信号に基づいて上記発声音声を音声認識する音声認識方法であって、所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換するステップと、上記変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換するステップと、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声を音声認識するステップと、上記音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力するステップとを含むことを特徴とする。

【0007】また、本発明に係る音声認識装置は、入力される発声音声文の音声信号に基づいて上記発声音声を音声認識する音声認識装置であって、所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変

(3)

特許第3027557号

5

6

換する第1の変換手段と、上記第1の変換手段によって変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換する第2の変換手段と、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記第2の変換手段によって変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声文を音声認識する音声認識手段と、上記音声認識手段によって音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力する単語抽出手段とを備えたことを特徴とする。

【0008】さらに、本発明に係る音声認識処理プログラムを記録した記録媒体は、入力される発声音声文の音声信号に基づいて上記発声音声文を音声認識する音声認識処理プログラムを記録した記録媒体であって、所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換するステップと、上記変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換するステップと、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声文を音声認識するステップと、上記音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力するステップとを含む音声認識処理プログラムを記録したことを特徴とする。

【0009】

【発明の実施の形態】以下、図面を参照して本発明に係る実施形態について説明する。

【0010】図1に本発明に係る実施形態の音声認識装置のブロック図を示す。本実施形態の音声認識装置においては、統計的言語モデルメモリ22に予め作成されて記憶されたN-gram又は可変長N-gramなどの統計的言語モデルを用いて絞り込んだ単語列の仮説を、文法的な制約によってさらに絞り込むと同時に、文法を逸脱した単語について、マーク付けすることによって、認識された発話中の信頼性の高い単語と低い単語が識別できるようにしたことを特徴とする。

【0011】図1において、単語照合部4は、公知のワーパス・ビタビ復号化法を用いて、入力される発声音声文の音声信号の特徴パラメータに基づいて上記発声音声文の単語仮説を検出し尤度を計算して出力し、次の

で、第1の単語仮説絞込部6は、単語照合部4からバッファメモリ5を介して出力される、終了時刻が等しく開始時刻が異なる同一の単語の単語仮説に対して、統計的言語モデルメモリ22内の統計的言語モデルを参照して、当該単語の先頭音素環境毎に、発声開始時刻から当該単語の終了時刻に至る計算された総尤度のうちの最も高い尤度を有する1つの単語仮説で代表させるように単語仮説の絞り込みを行う。ここで用いる統計的言語モデルは、学習用テキストデータに基づいて言語モデル生成部20により生成されたものであって、統計的言語モデル22は、品詞クラス間のバイグラム(N=2)を基本としたものであるが、単独で信頼できる単語は品詞クラスより分離させ、単独のクラスとして取り扱い、さらに、予測精度を向上させるため、頻出単語列に関してはそれらの単語を結合して一つのクラスとして取り扱い、長い単語連鎖の表現を可能にさせ、こうして、生成されたモデルは、品詞バイグラムと可変長単語N-gramとの特徴を併せ持つ統計的言語モデルとなり、遷移確率の精度と信頼性とのバランスをとられたものである。

【0012】本実施形態においては、文法的な制約として、文脈自由文法(CFG)で記述されたものを用いる。ただし、効率的な制約の適用を実現するため、文脈自由文法メモリ41に記憶された文脈自由文法(CFG)を、第1の文法変換部31によって予め有限状態オートマトン(FSA)に近似変換して、有限状態オートマトンメモリ42に格納する。さらに、多少の文法的な逸脱を許容するとともに、逸脱した単語にマーク付けするために、有限状態オートマトンメモリ42に格納された有限状態オートマトンを、第2の文法変換部32によって、単語の付加、削除及び置換の状態遷移を含む有限状態トランスデューサ(FST)に変換して有限状態トランスデューサ(FST)メモリ43に格納する。そして、第2の単語仮説絞込部7は、こうして作られた有限状態トランスデューサ(FST)を用いて、第1の単語仮説絞込部6から出力される単語列の仮説を絞り込むと同時に、文法を逸脱した単語にマーク付けする。最後に、単語抽出部8は、マーク付けされた単語(又は、さらにはその周辺の単語)を取り除くことにより、信頼性の高い単語列からなる発話断片を抽出して音声認識結果として出力する。

【0013】図1において、単語照合部4に接続され、音素HMMメモリ11に格納される音素HMMは、各状態を含んで表され、各状態はそれぞれ以下の情報を有する。

- (a) 状態番号
- (b) 受理可能なコンテキストクラス
- (c) 先行状態、及び後続状態のリスト
- (d) 出力確率密度分布のパラメータ
- (e) 自己遷移確率及び後続状態への遷移確率

なお、本実施形態において用いる音素HMMは、各分布

(4)

特許第3027557号

7

がどの話者に由来するかを特定するため、所定の話者混合HMMを変換して生成する。ここで、出力確率密度関数は3次元の対角共分散行列をもつ混合ガウス分布である。また、単語照合部4に接続され、単語辞書メモリ12に格納される単語辞書は、音素HMMの各単語毎にシンボルで表した読みを示すシンボル列を格納する。

【0014】図1において、話者の発声音声はマイクロホン1に入力されて音声信号に変換された後、特徴抽出部2に入力される。特徴抽出部2は、入力された音声信号を Λ/D 変換した後、例えばLPC分析を実行し、対数パワー、16次ケプストラム係数、 Δ 対数パワー及び16次 Δ ケプストラム係数を含む34次元の特徴パラメータを抽出する。抽出された特徴パラメータの時系列はバッファメモリ3を介して単語照合部4に入力される。

【0015】単語照合部4は、ワンパス・ビタビ復号化法を用いて、バッファメモリ3を介して入力される特徴パラメータのデータに基づいて、音素HMMと単語辞書とを用いて単語仮説を検出し尤度を計算して出力する。ここで、単語照合部4は、各時刻の各HMMの状態毎に、単語内の尤度と発声開始からの尤度を計算する。尤度は、単語の識別番号、単語の開始時刻、先行単語の違い毎に個別にもつ。また、計算処理量の削減のために、音素HMM及び単語辞書とに基づいて計算される総尤度のうちの低い尤度のグリッド仮説を削減する。単語照合部4は、その結果の単語仮説と尤度の情報を発声開始時刻からの時間情報（具体的には、例えばフレーム番号）とともにバッファメモリ5を介して単語仮説絞込部6に出力する。

【0016】単語仮説絞込部6は、単語照合部4からバッファメモリ5を介して出力される単語仮説に基づいて、統計的言語モデル22を参照して、終了時刻が等しく開始時刻が異なる同一の単語の単語仮説に対して、当該単語の先頭音素環境毎に、発声開始時刻から当該単語の終了時刻に至る計算された総尤度のうちの最も高い尤度を有する1つの単語仮説で代表させるように単語仮説の絞り込みを行った後、絞り込み後のすべての単語仮説の単語列のうち、最大の総尤度を有する仮説の単語列を認識結果として出力する。本実施形態においては、好ましくは、処理すべき当該単語の先頭音素環境とは、当該単語より先行する単語仮説の最終音素と、当該単語の単語仮説の最初の2つの音素とを含む3つの音素並びをいう。

【0017】例えば、図2に示すように、 $(i-1)$ 番目の単語 W_{i-1} の次に、音素列 a_1, a_2, \dots, a_n からなる i 番目の単語 W_i がくるときに、単語 W_{i-1} の単語仮説として6つの仮説 $W_a, W_b, W_c, W_d, W_e, W_f$ が存在している。ここで、前者3つの単語仮説 W_a, W_b, W_c の最終音素は $/x/$ であるとし、後者3つの単語仮説 W_d, W_e, W_f の最終音素は $/y/$ であるとす

8

る。終了時刻 t_e と先頭音素環境が等しい仮説（図2では先頭音素環境が $"x/a_1/a_2"$ である上から3つの単語仮説）のうち総尤度が最も高い仮説（例えば、図2において1番上の仮説）以外を削除する。なお、1から4番めの仮説は先頭音素環境が違いため、すなわち、先行する単語仮説の最終音素が x ではなく y であるので、1から4番めの仮説を削除しない。すなわち、先行する単語仮説の最終音素毎に1つのみ仮説を残す。図2の例では、最終音素 $/x/$ に対して1つの仮説を残し、最終音素 $/y/$ に対して1つの仮説を残す。

【0018】次いで、第1及び第2の文法変換部31、32の処理について詳述する。第1の文法変換部31は、文脈自由文法（CFG）から有限状態オートマトン（FSA）を自動生成する処理部である。文脈自由文法（CFG）は、次表に示す（ V_T, V_N, P, S ）の4つの組みで表される。

【0019】

【表1】

文脈自由文法（CFG）

 V_T ：終端記号の有限集合 V_N ：非終端記号の有限集合 P ：生成規則の有限集合 S ：開始記号

【0020】開始記号を生成規則を用いて書き換え、最終的に得られる終端記号列の集合が、この文脈自由文法で表される言語である。終端記号とは、生成規則の左辺に現れない記号であり、生成規則によって最終的に生成される記号列の構成要素である。非終端記号は、それ以外の生成規則に現れる記号である。生成規則は、一つの非終端記号を0個以上の終端又は非終端記号に書き換える規則である。次に、その規則の一例を示す。

【0021】

【数1】 $V_T = \{det, noun, pron, prep, verb\}$ $V_N = \{SENT, NP, VP, PP\}$ P ：次の表2参照。 $S = SENT$

【0022】ここで、 det は冠詞、 $noun$ は名詞、 $pron$ は代名詞、 $prep$ は前置詞、 $verb$ は動詞を示す。また、 $SENT$ は開始記号、 NP は名詞句、 VP は動詞句、 PP は前置詞句を示す。また、英語のための生成規則の一例を次の表に示す。

【0023】

【表2】

英語のための生成規則の一例

 $SENT \rightarrow NP, VP, NP$ $SENT \rightarrow SENT, PP$

(5)

特許第3027557号

9

10

NP → det, noun
 NP → pron
 NP → NP, PP
 PP → prep, NP
 VP → verb

*記号SENTは、NP, VP, 又はNPを生成すること
 を示している。本実施形態で用いる有限状態オートマ
 トン(FSA)は、入力記号列を受理するかしないかを決
 定する仮想的な機械であって、次に示す(Q, Σ,
 q₀, F, E)の5つの組みで表される。

【0025】

【表3】

【0024】ここで、例えば、表2の第1行目は、開始*

Q: 状態の有限集合

Σ: 入力記号の有限集合

q₀ ∈ Q: 初期状態

F ⊆ Q: 最終状態の有限集合

E ⊆ Q × (Σ ∪ {ε}) × Q:

状態遷移の有限集合Eは、遷移前の状態、入力記号、
 遷移後の状態の3つの組みからなる。

【0026】FSAの基本動作は、初期状態から始ま
 り、入力記号を読みつつ、状態を遷移するというもので
 ある。入力記号を読み終わったとき、最終状態に到達する
 ことが可能であれば、その記号列を受理する。状態遷移で
 使われているεは、特別な入力記号で、入力を読むこと
 なしに遷移が可能であることを表す。任意の状態におい
 て、ある入力記号が決まったとき、一意に遷移先の状態
 がきまるFSAを決定的であるという。任意のFSA
 は、決定的かつ状態数が最小である等価なFSAに
 一意に変換可能であるという性質を持っている。この変
 換処理は公知であって、例えば、従来技術文献「ジェイ
 ・ホップクロフトほか(訳者野崎ほか)、“オートマトン
 言語理論—計算論I”、pp. 17-70及びpp. 71
 -98、サイエンス社、昭和59年8月25日発行」な
 どに開示されている。しかしながら、FSAで表現でき
 る言語のクラスよりも、CFGで表現できる言語のクラ
 スが大きい。一般的には、CFGを等価なFSAに
 自動変換することはできない。従って、本実施形態で
 は、例えば、ペレイラのアルゴリズム(例えば、前掲の
 従来技術文献「F. C. N. Pereira, et al.」)を用い
 て、CFGを近似的に、FSAに変換する。すなわち、
 本実施形態では、ペレイラのアルゴリズムに基づいて、
 第1の文法変換処理を実行して、CFGを近似されたF
 SAに変換する。第1の文法変換処理は、具体的には、
 表4のCFG規則を例にとると、次のようにして行われ
 る。

【表4】

S → a, S, b.
 S → c.

この文法では、非終端記号はSのみで、開始記号を兼ね
 る。a, bが終端記号である。

(1) 特別な非終端記号S'及び特別な規則「S' → 開

始記号」すなわち「S' → S」を導入する。

(2) CFG規則の右辺の終端/非終端記号の前後にド
 ットを記した、ドットつき規則を導入する。ドットは、
 入力終端記号列に対する現在の文法規則上の処理位置
 を表す。

(3) {S' →, S}の閉包を求める。閉包とは、ドッ
 トの直後に現れる非終端記号を左辺に持つすべての規則
 について、右辺の先頭にドットのついたドットつき規則
 を再帰的に追加してゆく処理である。この結果、ドッ
 付き規則の集合{S' →, S, S →, a S b, S →,
 c}が得られる。

(4) 次に示す手順によって、ドットつき規則の閉包集
 合を状態としたFSAを構築する。図12乃至図15の
 四角又は丸はFSAの状態を表し、矢印は遷移を表す。
 二重線の四角又は二重丸は、最終状態を表す。また、始
 端をもたない矢印の終端となっている状態は、初期状態
 を表す。矢印上に記された記号は、入力記号(CFGの
 終端記号に対応する。)を表す。

(4a) 閉包集合{S' →, S, S →, a S b, S →,
 c}を初期状態にする。

(4b) 初期状態となるドットつき規則の閉包集合から
 始まり、そこから終端・非終端記号で遷移できるドッ
 付き規則の閉包集合を再帰的に生成する(図12参
 照。)。ドットつき規則「S' → S,」を含む状態を、
 最終状態とする。

(4c) 右辺の最終端にドットのある規則と、その規則
 の左辺の終端記号の直後にドットがあるようなドットつ
 き規則をすべての閉包集合から探す。前者の規則を含む
 閉包集合から、後者の規則を含む閉包集合へ、ε遷移を
 追加する(図13参照。)

(4d) 非終端記号のついた遷移を消去する(図14参
 照。)

(5) さらに、図15に示すように、決定的かつ最小な
 FSAに変換する。

(6)

特許第3027557号

II

12

【0027】図3に、数1のCFGの例に対して第1の文法変換処理を実行することにより、作成した近似されたFSAを示す。丸は状態を、矢印は遷移を表す。二重丸は最終状態をあらわす。また、始端をもたない矢印の終端となっている状態は、初期状態を表す。矢印上に記された記号は、入力記号を表す。当該FSAは、決定的であり、最小である。

【0028】先の例のCFGの終端記号、すなわちFSAの入力記号は、単語の品詞を表している。“I saw a girl with a telescope”という例文は、
 “I (pron) saw (verb) a (det) girl (noun) with (prep) a (det) telescope (noun)”
 のように、単語と品詞を対応づけることができる。CFGの生成規則においては、“with (prep) a (det) telescope (noun)”という前置詞句(PP)は、規則「NP→NP, PP」を用い *

*て、生成されるか、それとも規則「SENT→SENT, PP」を用いて生成されるか曖昧である。それに対して、変換されたFSAにおいては、状態遷移の曖昧性なく、決定的にこの例文を受理することが可能である。最適化されたFSAのもつ、このような性質により、CFGをそのまま文法的制約の表現として用いた場合と比べて、より効率的な制約適用が可能となる。

【0029】第2の文法変換部32は、文法的制約を表現するFSAを、多少の文法的な逸脱を許容するとともに、逸脱した単語にマーク付する有限状態トランスデューサ(FST)に変換する処理部である。FSTは、FSAに出力記号を追加した次表の6つの組(Q, Σ, Σ', q₀, F, E')で表されるもので、入力シンボル列が受理されるとき、対応する出力シンボル列を出力する。

【0030】

【表5】

Q: 状態の有限集合

Σ: 入力記号の有限集合

Σ': 出力記号の有限集合

q₀ ∈ Q: 初期状態

F ⊆ Q: 最終状態の有限集合

E' ⊆ Q × (Σ ∪ {ε}) × Σ' × Q:

状態遷移の有限集合E'は、遷移前の状態、入力記号、出力記号列、遷移後の状態の4つの組からなる。

【0031】第2の文法変換部32では、決定的なFSAから、図5乃至図7の第2の文法変換処理によって、付加・脱落・置換を考慮することで、文法的な逸脱を許容するFSTを生成する。

【0032】図5乃至図7は、図1の第2の文法変換部32によって実行される第2の文法変換処理を示すフローチャートである。図5において、まず、ステップS1において、有限状態オートマトンメモリ42から有限状態オートマトン(FSA)を読み込み、ステップS2で初期設定処理を実行するため、状態遷移の有限集合E'にゼロのデータφを代入するとともに、出力記号の有限集合Σ'に入力記号の有限集合Σを代入する。次いで、ステップS3でa ∈ Σの各要素aに対して付加記号Ins<a>、脱落記号Del<a>及び置換記号Subst<a>を出力記号の有限集合Σ'に追加する。そして、ステップS4でa ∈ Σのすべての要素aについてステップS3の処理を実行したか否かが判断され、NOのときはステップS3の処理を繰り返す一方、YESのときはステップS5に進む。ステップS5でFSAの各状態遷移(q₁, a, q₂)に対して状態遷移に有限集合E'に状態遷移を追加する。そして、ステップS6でFSAのすべての状態遷移(q₁, a, q₂)に対してステップS5の処理を実行したか否かが判断され、NOのと

きはステップS5の処理を繰り返す一方、YESのときは図6のステップS7に進む。

【0033】図6のステップS7において、1つの状態遷移(q₁, a, a, q₂)に対して脱落を示す状態遷移(q₁, ε, Del<a>, q₂)を追加する。次いで、ステップS8でb ∈ Σの各要素bに対して置換を表す状態遷移(q₁, a, Subst, q₂)を追加する。そして、ステップS9でb ∈ Σのすべての要素bについてS8の処理を実行したか否かが判断され、NOのときはステップS8の処理を繰り返す一方、YESのときはステップS10に進む。ステップS10ですべての状態遷移(q₁, a, a, q₂)に対してS7の処理を実行したか否かが判断され、NOのときはステップS7以降の処理を繰り返す一方、YESのときはステップS11に進む。次いで、ステップS11でq ∈ Qの1つの要素qに対して、かつa ∈ Σの各要素aに対して付加を表す状態遷移(q, a, Ins<a>, q)を追加する。そして、ステップS12でa ∈ Σのすべての要素aに対してS11の処理を実行したか否かが判断され、NOのときはステップS11の処理を繰り返す一方、YESのときはステップS13に進む。さらに、ステップS13でq ∈ Qのすべての要素qに対してS11の処理を実行したか否かが判断され、NOのときはステップS1

(7)

特許第3027557号

13

1の処理を繰り返す。一方、YESのときは図7のステップS14に進む。最後に、図7のステップS14で得られた有限状態トランスデューサ(FST)を有限状態トランスデューサ(FST)メモリ43に書き込む。これで、第2の文法変換処理を終了する。

【0034】図4に、図3の近似されたFSAから生成したFSTを示す。スラッシュの左辺が入力記号、右辺が出力記号を表す。疑問符が含まれている状態遷移は、疑問符をそれぞれの入力シンボルに置き換えた複数の状態遷移に相当する。

【0035】第2の単語仮説生成部7は、公知の単語グラフ(例えば、従来技術文献「T. Shimizu, et al., "Spontaneous dialogue speech recognition using cross-word context constrained word graphs", In Proceedings of ICASSP, 1996年」参照。)の形式で表現された単語列仮説のもともらしさを、第2の文法変換部32で生成したFSTを用いて、再尤度づけするとともに、単語とFSTの出力記号とを対応づける処理部である。単語グラフは、入力記号が単語であるFSAを状態遷移毎に遷移尤度をもつように拡張したものとみなすことができる。また、この第2の単語仮説生成部7では、あらかじめFSTの方にも、状態遷移の尤もらしさを表現する遷移尤度が与えられていると想定する。FSTの遷移尤度の与え方には、さまざまな方法が考えられるが、例えば、付加・脱落・挿入を表す遷移を通ったときに、ペ

14

ナルティを課す値として-1を、その他の遷移には、ペナルティなしを表す0を与える。さらに、単語グラフの遷移尤度XとFSTの遷移尤度Yから新たな尤度を求める関数fを用意する。この関数fについても、様々な実現が考えられるが、例えば単語グラフの尤度を無視して、 $f(X, Y) = Y$ のような関数を用いる。

【0036】単語グラフと尤度つきFSTの両方に受理される単語列について、この再尤度づけ関数fを用いて計算する累積尤度を最大化するように、「単語とFSTの出力記号の対応」列を求めるのが、第2の単語仮説生成部7である。第2の単語仮説生成部7によって実行された第2の単語仮説生成処理のフローチャートを図8乃至図10に示す。この処理のアルゴリズムは、例えば従来技術文献「ジェイ・ホップクロフトほか、「オートマトン理論—計算論I」、サイエンス社、pp. 77」に開示され、複数のFSAが与えられたとき、これらすべてのFSAに共通して受理される入力記号列を受理するようなFSAを求めるアルゴリズムと、例えば従来技術文献「J. Pearl, "Heuristics", Addison-Wesley, Reading, MA, 1984年」に開示され、各辺にコストの定義された有向グラフにおいて、最適パスを効率良く求めるA*アルゴリズムを融合したものである。当該処理中で用いられている変数などの意味を次の表に示す。

【0037】

【表6】

単語グラフ: $(Q_w, \Sigma_w, q_0, F_w, E_w)$

Q_w : 状態の有限集合

Σ_w : 入力記号(単語)の有限集合

$q_0 \in Q_w$: 初期状態

$F_w \subseteq Q_w$: 最終状態の有限集合

$E_w \subseteq Q_w \times (\Sigma_w \cup \{\epsilon\}) \times Q_w$: 状態遷移の有限集合。

FST: $(Q_{fst}, \Sigma_{fst}, \Sigma'_{fst}, p_0, F_{fst}, E'_{fst})$

Q_{fst} : 状態の有限集合

Σ_{fst} : 入力記号(品詞)の有限集合

Σ'_{fst} : 出力記号の有限集合

$p_0 \in Q_{fst}$: 初期状態

$F_{fst} \subseteq Q_{fst}$: 最終状態の有限集合

$E'_{fst} \subseteq Q_{fst} \times (\Sigma_{fst} \cup \{\epsilon\}) \times \Sigma'_{fst} \times Q_{fst}$: 状態遷移の有限集合。

Score_w, Score_{fst}: 状態遷移に遷移尤度を与える関数

$g[q, p]$:

初期状態(q_0, p_0)から(q, p)までの最尤パスを格納する添字つき変数。

本変数に格納されるパスとは、単語グラフの状態とFSTの状態の組を新たな状態としたときの状態遷移の列をいう。この新たな状態遷移は、(単語グラフの状態とFSTの状態の組、単語、品詞、FSTの出力記号列、単語グラフの状態とFSTの状態の組)の4つ組からなる。

【0038】次いで、図8乃至図10を参照して、第2の単語仮説絞込処理について説明する。図8において、まず、ステップS21で初期設定処理を実行し、ここで、単語グラフの状態とFSTの状態の組の集合OPENに初期状態 $\{(q_0, p_0)\}$ を代入し、初期状態 $[q_0, p_0]$ までの最尤パス $g[q_0, p_0]$ に長さ0のパスを表すNULLを代入する。次いで、ステップS22で状態 $(q_1, q_2) \in \text{OPEN}$ のうちパス $g[q_1, q_2]$ の尤度が最大のものを1つ取り出す。そして、ステップS23で $q_1 \in E_{\text{seg}}$ かつ $q_2 \in E_{\text{fst}}$ であるか否かを判断し、YESであれば、ステップS24で $g[q_1, q_2]$ を最尤パスと判断して単語抽出部8に出力して当該第2の単語仮説絞込処理を終了する。一方、ステップS23でNOであるときは、ステップS25で状態 (q_1, q_2) を変数データOPENから取り除き、図9のステップS26に進む。

【0039】図9のステップS26で、状態 $(q_2, \epsilon, \alpha, q_4) \in E_{\text{fst}}$ である状態 q_4 について、

(a) 変数データOPENに状態 (q_1, q_2) を追加する。

(b) パス $g[q_1, q_2]$ の後ろに $((q_1, q_2, \epsilon, \epsilon, \alpha, (q_1, q_4)))$ を継ぎ足したものを変数 $g[q_1, q_4]$ に代入する。

(c) パス $g[q_1, q_2]$ の尤度として、 $g[q_1, q_2]$ の尤度 $+f(0, \text{Score}_{\text{fst}}(q_2, \epsilon, \alpha, q_4))$ を設定する。

そして、ステップS27で状態 $(q_2, \epsilon, \alpha, q_4) \in E_{\text{fst}}$ であるすべての状態 q_4 についてS26の処理を実行したか否かが判断され、NOのときはステップS26の処理を繰り返す一方、YESのときはステップS28に進む。次いで、ステップS28では、 $(q_1, w, q_3) \in E_{\text{seg}}$ かつ $(q_2, \text{pos}, \alpha, q_4) \in E_{\text{fst}}$ かつ $w \in \text{pos}$ である状態 (q_3, q_4) を見つけて、

(a) 変数データOPENに状態 (q_3, q_4) を追加する。

(b) パス $g[q_1, q_2]$ の後ろに、 $((q_1, q_2, w, \text{pos}, \alpha, (q_3, q_4)))$ を継ぎ足したものを変数 $g[q_3, q_4]$ に代入する。

(c) パス $g[q_3, q_4]$ の尤度として、 $g[q_1, q_2]$ の尤度 $+f(\text{Score}_{\text{seg}}(q_1, w, q_3), \text{Score}_{\text{fst}}(q_2, w, \alpha, q_4))$ を設定する。

そして、ステップS29で、ステップS28の条件のすべての状態 (q_3, q_4) に対してS28の処理を実行したか否かが判断され、NOのときはステップS28の処理を繰り返す一方、YESのときは図10のステップS30に進む。

【0040】ステップS30では、最尤パス候補の終端の状態集合を示す変数データOPENが空集合であるか否かが判断され、NOのときはステップS22に戻り、

記の処理を繰り返す一方、YESのときは、ステップS31で「単語列の仮説なし」を単語抽出部8に出力し、当該第2の単語仮説絞込処理を終了する。

【0041】例えば、単語列“hi saw girl with a telescope”は単語グラフの特殊な場合と考えることができるが、各々の単語の品詞が次の(a)のようであるとすると、第2の単語仮説絞込処理における最尤パス探索によって、単語と Σ^{fst} *の要素は、次の(b)のように対応づけることができる。

(a) hi (interj), saw (verb), girl (noun), with (prep), a (det), telescope (noun)

(b) hi (Subst<pron>), saw (verb), ϵ (Del<det>), girl (noun), with (prep), a (det), telescope (noun)

【0042】最後に、単語抽出部8は、信頼性の高い単語を抽出する処理を実行する処理部であって、第2の単語仮説絞込部7で得られた結果の最尤パスについて、次の方法で信頼性の低いと思われる単語を削除する。すなわち、Subst, Del, Insでマーク付された単語を削除する。この変形例としては、Subst, Del, Insでマーク付された単語および、それに隣接する単語を削除してもよい。例えば、“hi (Subst<pron>), saw (verb), ϵ (Del<det>), girl (noun), with (prep), a (det), telescope (noun)”の例では、本実施形態の方法によって、“saw (verb)”, “girl (noun), with (prep), a (det), telescope (noun)”が、変形例の方法によって、“with (prep), a (det), telescope (noun)”が得られる。

【0043】以上の実施形態においては、当該単語の先頭音素環境とは、当該単語より先行する単語仮説の最終音素と、当該単語の単語仮説の最初の2つの音素とを含む3つの音素並びとして定義されているが、本発明はこれに限らず、先行する単語仮説の最終音素と、最終音素と連続する先行する単語仮説の少なくとも1つの音素とを含む先行単語仮説の音素列と、当該単語の単語仮説の最初の音素を含む音素列とを含む音素並びとしてもよい。

【0044】以上の実施形態において、特徴抽出部2と、単語照合部4と、第1と第2の単語仮説絞込部6, 7と、言語モデル生成部20と、第1と第2の文法変換部31, 32とは、例えば、デジタル電子計算機で構成される一方、バッファメモリ3, 5と、音素HMMメモリ11と、単語辞書メモリ12と、学習用テキストデー

(9)

特許第3027557号

17

タメモリ21と、統計的言語モデルメモリ22と、文脈自由文法メモリ41と、有限状態オートマトンメモリ42と、有限状態トランスデューサメモリ43とは、例えばハードディスクメモリなどの記憶装置で構成される。なお、以上の実施形態において、単語照合部4と、第1と第2の単語仮説絞込部6、7は音声認識手段を構成している。

【0045】本実施形態の音声認識装置のすべての音声認識処理のための音声認識処理プログラムを、光ディスク又は光磁気ディスクなどの記録媒体に記録して、例えば1つのデジタル電子計算機で当該処理プログラムを実行するように構成してもよい。

【0046】

【実施例】本発明者は、本実施形態の音声認識装置の性能を評価するために、以下の実験を行った。ここでは、N-gramに基づく統計学的言語モデルのみを使用して認識される最上位（トップベスト）単語列の信頼性と、最上位の単語列をロバスト・パーズング、すなわち頑健なパーズング（本実施形態における第2の単語仮説絞込部7及び単語抽出部8によって構成されるパーズング処理部によるパーズング処理をいう。）することにより取得される断片的な単語列の信頼性を比較した。音声認識タスクには、本特許出願人が所有する自然音声データベースに含まれる55のホテル予約会話を使用した。このデータベースの場合、会話は2カ国語であり、話者は通訳を介して会話している。認識実験には、のべ22,695単語からなる日本語による1,535発話を使用した。また、文脈自由文法は音声認識用に開発されたものを使用した。この文法は1832項目の規則で構成され、文ではなくポーズ挿入可能な発話断片を文法の単位としている。文法は認識実験に使用された55会話に含まれる9会話を使用して開発した。N-gramに基づく統計学的言語モデルとしては、先の55会話を含む98会話で作成した可変長N-gramを使用した。

【0047】単語セグメントの信頼性の評価には、情報検索の研究分野で使用される適合率を使用する。適合率は以下のように定義される。

【数2】

適合率＝合致した単語数／認識された単語数×100
合致した単語数とは認識された単語列と正解単語列との間で一致する単語の最大数である。一般的な認識率とは分母が異なっており、分母が正解単語数ならば通常の認識率となる。

【0048】可変長N-gramを使用した最上位（トップベスト）の認識結果における適合率は、68%であった。これに対して、頑健なパーズングによって得られた発話断片の適合率は73%であった。この実験は、我々が提案する音声認識方法を使用すれば、発話に含まれる信頼性の高い断片を獲得することが可能であることを示している。

18

【0049】音声認識のためには、取得する部分発話断片の信頼性だけでなく頑健なパーズングの正解単語カバー率を高めることも必要である。

【数3】正解単語の適用範囲＝（頑健なパーズングの出力正解単語数）／（頑健なパーズングの入力正解単語数）×100

通常、同じ文法的制約を用いる場合であっても、本実施形態の頑健なパーズングの柔軟性によって得られる発話断片の信頼性と正解単語のカバー率との間にはトレードオフが存在する。このトレードオフの関係を調べるため、本発明に係る本実施形態のきつい頑健なパーズングについても認識実験を行った。

【0050】きつい頑健なパーズングは、挿入、削除及び置換を隣接単語を含めて無視することによって達成される。これは隣接単語が、挿入、削除、置換の影響を受けて信頼性が低いと考えられるためである。例えば、このきつい頑健なパーズング法を使用した場合、“hi (Subst<pron>), saw (verb), r (Del<det>), girl (noun), with (prep), a (det), telescope (noun)”からは、“saw (verb)”, “girl (noun), with (prep), a (det), telescope (noun)”が得られる。

【0051】図11は、頑健なパーズングによって棄却された単語の割合を示している。図11から明らかなように、本実施形態のきつい頑健なパーズングを使用すれば、73%の信頼性を81%に向上させることができる。しかし、反対に正解単語のカバー率は89%＝（47%＋14%）／68%から69%＝47%／68%に低減する。現在使用している文法の最大の問題は、その辞書項目が実験に用いた認識対象発話を不十分にカバーしていないという点にある。我々の頑健なパーズング法では、文法の語彙項目に含まれない単語は挿入または置換と見なされる。この語彙数の欠如を反映して、我々の文法の認識対象発話カバー率はゆるい頑健なパーズング法を用いた場合89%であり、きつい方法を用いた場合は71%である。これらの比率はそれぞれ、上述の正解単語カバー率89%と69%にほぼ等しい。従って、文法上の辞書項目が十分でさえあれば、正解単語をさらにカバーすることが可能となる。

【0052】頑健な音声対話システムを達成するため、我々は、N-gramに基づく統計学的言語モデルを使用した認識結果を頑健なパーズング処理を行うことにより信頼できる部分的な発話断片を得る認識方法を発明した。我々の方法は、CFGによって表現された文法的制約を効果的に適用可能な表現に近似変換することによって頑健なパーズングに使用している。後段の文法を文法的制約として使用することにより、音声対話システム全体の性能を向上させることができる。自然発話音声認識の実験を通して、我々の方法がN-gramに基づく統

(1 0)

特許第3027557号

19

20

計学的言語モデルだけを使用する従来の連続音声認識と比べて信頼性の高い部分発話断片を取得可能であることを示した。信頼できる発話断片を求めた後、その情報を用いて、それ以外の部分を求めなおす再探索法にも拡張が可能である。我々の認識方法は、マルチパス探索法に基づく頑健な認識にも拡張可能である。こうした方法は最初のパスの後に信頼できるセグメントの情報を使用する。特に未知語を含んだ発話の認識には、このタイプの探索方法が不可欠である。

【0053】以上説明したように、本実施形態によれば、以下の特有の効果を有する。

(a) 統計的な言語制約と、文法的な制約とを、本実施形態のように組み合わせることにより、単独の制約を用いた場合より、文法的でない自然発話を、高い精度で認識可能であり、信頼性の高い区間を見つけることができる。

(b) 有限状態オートマトンに基づいた近似的な文法制約を用いることで、効率的な制約適用が可能である。従って、音声認識処理を従来技術に比較して高速で実行することができる。

(c) 文法を逸脱した部分の単語を無視することによって、信頼性高く認識された発話断片を求めることができる。これによって、求められた発話断片に対する音声認識率を従来技術に比較して大幅に向上させることができる。

(d) 本実施形態は、信頼性高く認識された発話断片に基づいて、発話理解処理を進める音声対話方法及び装置や、この発話断片情報をもとに、認識候補の再探索を行うマルチパス探索方法に適用することができる。これらの方法及び装置では、それぞれ頑健な音声理解、音声認識が可能となる。

【0054】

【発明の効果】以上詳述したように本発明によれば、入力される発声音声文の音声信号に基づいて上記発声音声文を音声認識する音声認識方法及び装置、並びに、音声認識処理プログラムを記録した記録媒体であって、所定の文脈自由文法を、文法的制約を表わす有限状態オートマトンに変換し、上記変換された有限状態オートマトンを、自然発話における単語の挿入、脱落、置換の状態遷移を含み文法的制約を表わす有限状態トランスデューサに変換し、入力される発声音声文の音声信号に基づいて、所定の統計的言語モデルと、上記変換された有限状態トランスデューサとを参照して、自然発話における単語の挿入、脱落、置換に該当する単語にマーク付けするように上記発声音声文を音声認識し、上記音声認識された結果に基づいて、上記マーク付けされた単語、もしくは、上記マーク付けされた単語及びそれに隣接する単語を、文法を逸脱した部分の単語として削除して音声認識結果を出力する。

【0055】従って、本発明によれば、以下の特有の効

果を有する。

(a) 統計的な言語制約と、文法的な制約とを、本発明のように組み合わせることにより、単独の制約を用いた場合より、文法的でない自然発話を、高い精度で認識可能であり、信頼性の高い区間を見つけることができる。

(b) 有限状態オートマトンに基づいた近似的な文法制約を用いることで、効率的な制約適用が可能である。従って、音声認識処理を従来技術に比較して高速で実行することができる。

(c) 文法を逸脱した部分の単語を無視することによって、信頼性高く認識された発話断片を求めることができる。これによって、音声認識率を従来技術に比較して大幅に向上させることができる。

【図面の簡単な説明】

【図1】 本発明に係る実施形態である音声認識装置のブロック図である。

【図2】 図1の音声認識装置における第1の単語仮説絞込部6の処理を示すタイミングチャートである。

【図3】 図1の第1の文法変換部31によって変換された近似された有限状態オートマトン(FSA)の一例を示す図である。

【図4】 図1の第2の文法変換部32によって変換された近似された有限状態トランスデューサ(FST)の一例を示す図である。

【図5】 図1の第2の文法変換部32によって実行される第2の文法変換処理の第1の部分を示すフローチャートである。

【図6】 図1の第2の文法変換部32によって実行される第2の文法変換処理の第2の部分を示すフローチャートである。

【図7】 図1の第2の文法変換部32によって実行される第2の文法変換処理の第3の部分を示すフローチャートである。

【図8】 図1の第2の単語仮説絞込部7によって実行される第2の単語仮説絞込処理の第1の部分を示すフローチャートである。

【図9】 図1の第2の単語仮説絞込部7によって実行される第2の単語仮説絞込処理の第2の部分を示すフローチャートである。

【図10】 図1の第2の単語仮説絞込部7によって実行される第2の単語仮説絞込処理の第3の部分を示すフローチャートである。

【図11】 図1の音声認識装置のシミュレーション結果であって、第2の単語仮説絞込部7によって棄却された単語の割合を示す図である。

【図12】 図1の第1の文法変換部31によって実行される第1の文法変換処理におけるドットつき規則の閉包集合と集合間の遷移を示す状態遷移図である。

【図13】 図1の第1の文法変換部31によって実行される第1の文法変換処理における遷移の追加を示す

(1 1)

特許第3027557号

21

22

状態遷移図である。

【図14】 図1の第1の文法変換部31によって実行される第1の文法変換処理における非終端記号の遷移の削除を示す状態遷移図である。

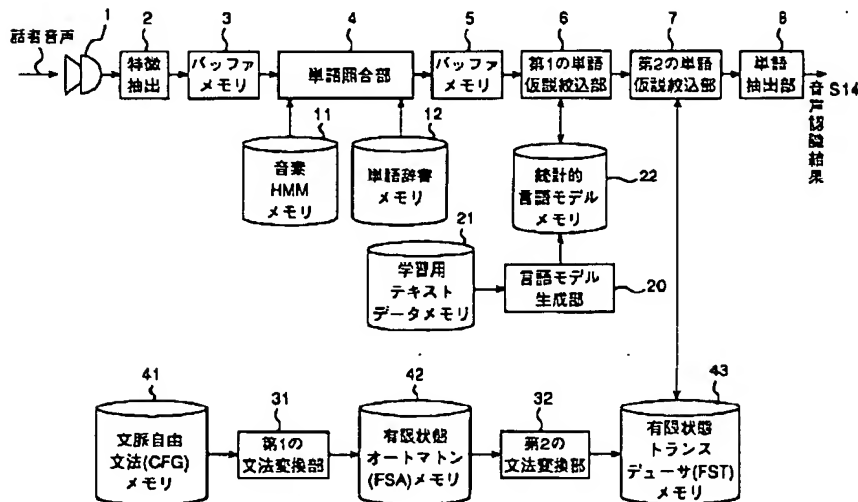
【図15】 図1の第1の文法変換部31によって実行される第1の文法変換処理における決定的かつ最小なFSAへの変換を示す状態遷移図である。

【符号の説明】

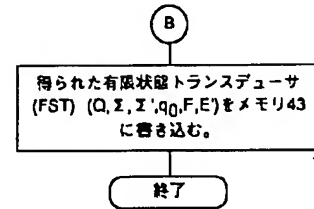
- 1…マイクロホン、
2…特徴抽出部、
3…バッファメモリ、
4…単語照合部、
5…バッファメモリ、
6…第1の単語仮説生成部、
7…第2の単語仮説生成部、
8…単語抽出部、
11…音素HMMメモリ、
12…単語辞書メモリ、
20…言語モデル生成部、
21…学習用テキストデータメモリ、
22…統計的言語モデルメモリ、
31…第1の文法変換部、
32…第2の文法変換部、
41…文脈自由文法(CFG)メモリ、
42…有限状態オートマトン(FSA)メモリ、
43…有限状態トランスデューサ(FST)メモリ、

- 7…第2の単語仮説生成部、
8…単語抽出部、
11…音素HMMメモリ、
12…単語辞書メモリ、
20…言語モデル生成部、
21…学習用テキストデータ、
22…統計的言語モデル、
31…第1の文法変換部、
32…第2の文法変換部、
41…文脈自由文法(CFG)メモリ、
42…有限状態オートマトン(FSA)メモリ、
43…有限状態トランスデューサ(FST)メモリ、

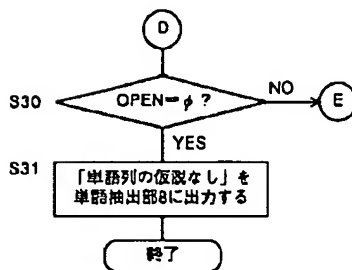
【図1】



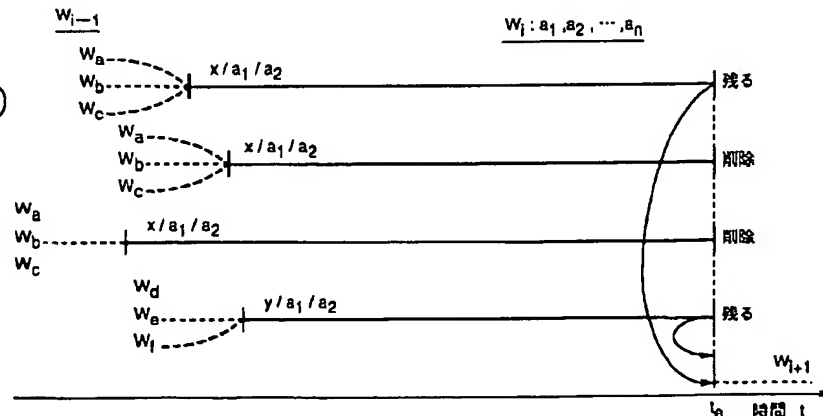
【図7】



【図10】



【図2】

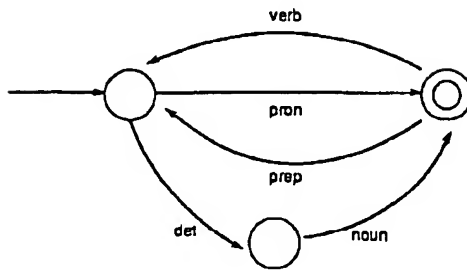


(1 2)

特許第3027557号

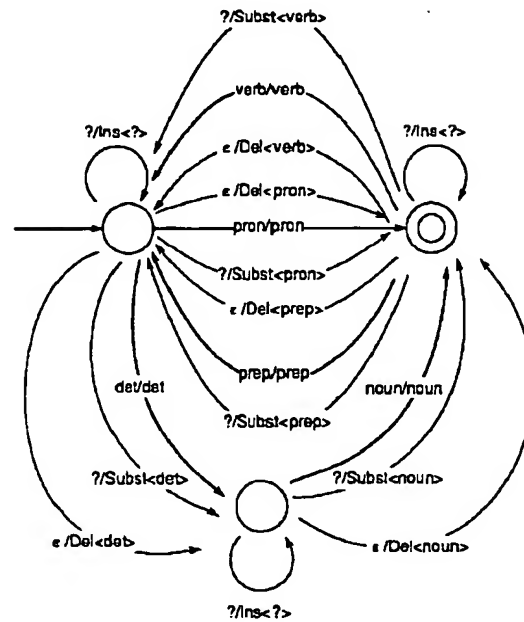
【図3】

近似された有限状態オートマトン(FSA)



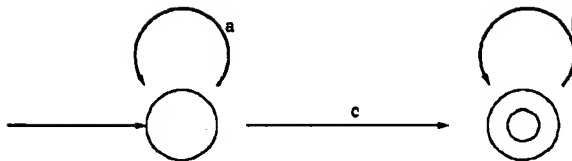
【図4】

生成された有限状態トランスデューサ(FST)



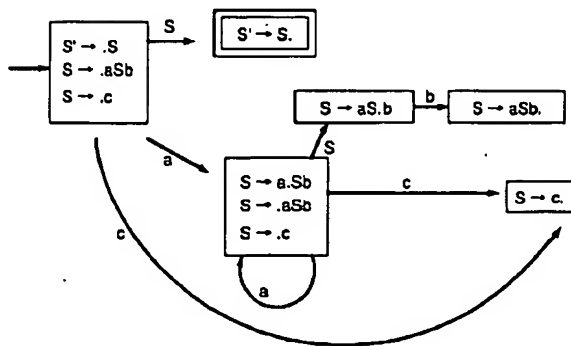
【図15】

決定的かつ最小なFSAに変換



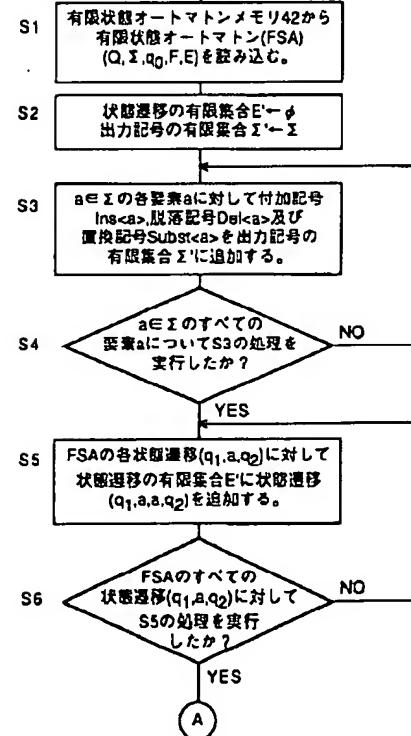
【図12】

ドットつき規則の閉包集合と集合間の遷移



【図5】

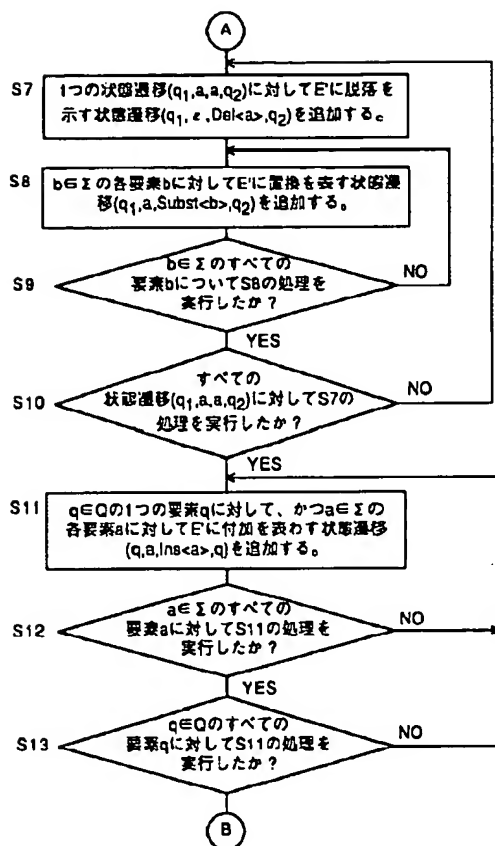
第2の文法変換処理



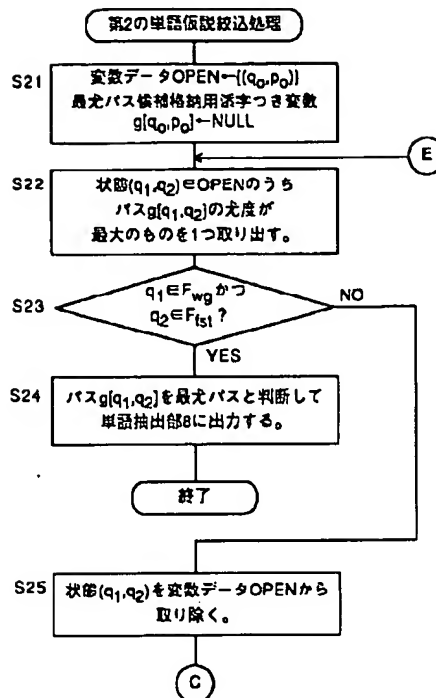
(13)

特許第3027557号

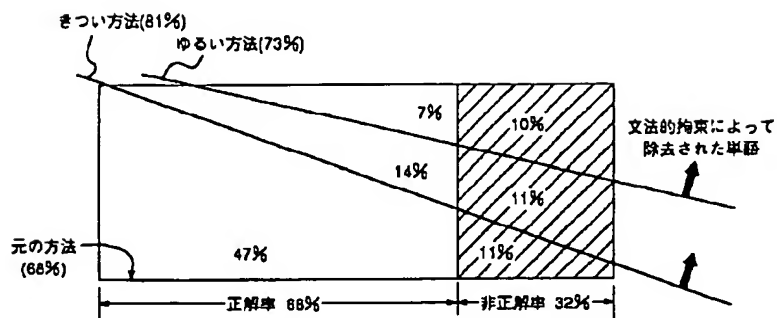
【図6】



【図8】



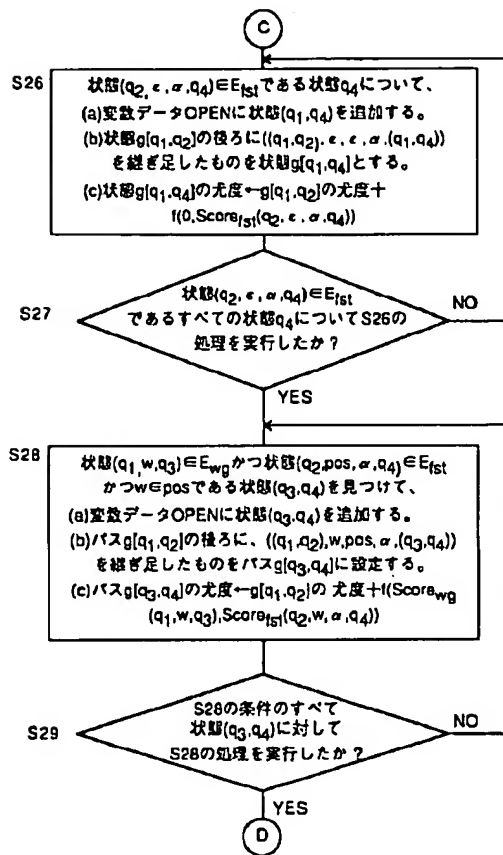
【図11】



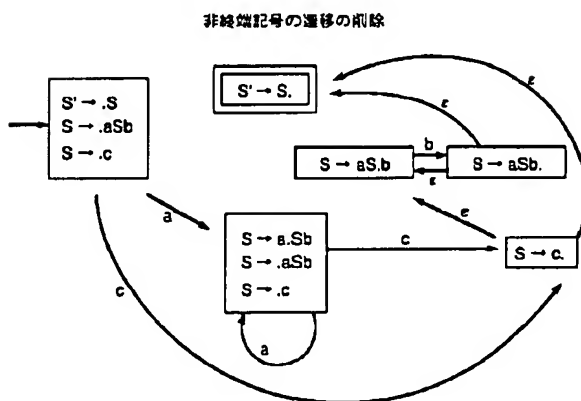
(14)

特許第3027557号

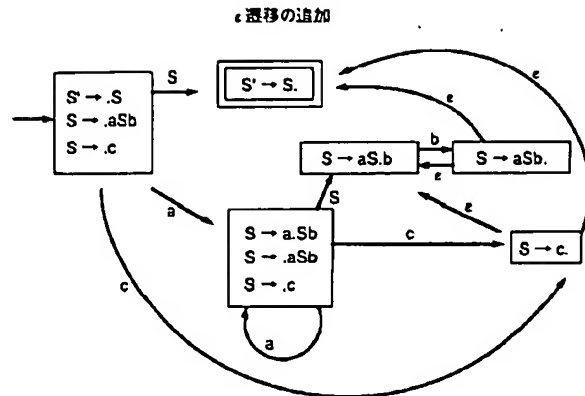
【図9】



【図14】



【図13】



(1 5)

特許第3027557号

フロントページの続き

(72)発明者 匂坂 芳典

京都府相楽郡精華町大字乾谷小字三平谷
5番地 株式会社エイ・ティ・アール音
声翻訳通信研究所内

(56)参考文献 特開 平8-123476 (J P, A)

(58)調査した分野 (Int.Cl.⁷, DB名)

G10L 15/00 - 15/28

J I C S Tファイル (J O I S)